Canad. Math. Bull. Vol. **24** (3), 1981

# CONSTRUCTION OF SATISFACTION CLASSES FOR NONSTANDARD MODELS

BY

H. KOTLARSKI, S. KRAJEWSKI AND A. H. LACHLAN

ABSTRACT. Given a resplendent model $\mathcal{M}$ for Peano arithmetic there exists a full satisfaction class over $\mathcal{M}$, i.e. an assignment of truth-values, to all closed formulas in the sense of $\mathcal{M}$ with parameters from $\mathcal{M}$, which satisfies the usual semantic rules. The construction is based on the consistency of an appropriate system of $\mathcal{M}$-logic which is proved by an analysis of standard approximations of nonstandard formulas.

This paper is a contribution to the study of the concept of satisfaction of nonstandard formulas, i.e. formulas in the sense of some nonstandard model of arithmetic.

Let PA (Peano arithmetic) be formulated in a finite language without function symbols and with logical operators $\neg$, $\vee$, and $\exists$. (Below we use an analysis of the structure of formulas which would become unmanageable were function symbols allowed.) Let $L$ denote the chosen language. As with other notations for languages, $L$ will also denote the set of formulas of the language. Let $\mathcal{M}$ be a model of PA then $L(\mathcal{M})$ is the language obtained by adjoining constants naming the elements of $\mathcal{M}$, and $^*L(\mathcal{M})$ denotes the set of all formulas of $L(\mathcal{M})$ in the sense of $\mathcal{M}$. By a *full satisfaction class* for $\mathcal{M}$ we mean a subset $\Sigma$ of the sentences of $^*L(\mathcal{M})$ containing all true atomic and negated atomic sentences satisfying the following three conditions:

(i) for every sentence $\varphi$ of $^*L(\mathcal{M})$ exactly one of $\varphi$ and $\neg\varphi$ belongs to $\Sigma$

(ii) for all sentences $\varphi$, $\psi$ of $^*L(\mathcal{M})$, $\varphi \vee \psi \in \Sigma$ if and only if at least one of $\varphi$ and $\psi$ is in $\Sigma$

(iii) for each sentence $\exists x\varphi(x)$ of $^*L(\mathcal{M})$, $\exists x\varphi(x) \in \Sigma$ if and only if $\varphi(a) \in \Sigma$ for some $a \in |\mathcal{M}|$.

Let $\mathcal{U}$ be a structure in which the standard model $\mathcal{N}$ of PA is definable such that $Th(\mathcal{N})$, the set of true sentences of $L(\mathcal{N})$, or, strictly speaking, the set of Gödel numbers of such sentences, is definable in $\mathcal{U}$. Let $^*\mathcal{U}$ be an elementary extension of $\mathcal{U}$. Robinson [8] observed that membership in $^*Th(\mathcal{N})$ is a possible definition of truth for sentences of $^*L(\mathcal{N})$. He called this the "internal" definition of truth and contrasted it with an "external" definition formulated in terms of Skolem functions. The latter definition is not applicable to all

Received by the editors June 19, 1979 and in revised form February 21, 1980.

sentences of $^*L(\mathcal{N})$ and Robinson showed that even where it applies it is not in general equivalent to the internal definition. In our terminology $^*Th(\mathcal{N})$ is a full satisfaction class, but Robinson's external truth seems to have no immediate relevance to the present work.

Our main result is that any resplendent model $\mathcal{M}$ of PA has a full satisfaction class. We will also give a necessary and sufficient condition for the existence of a full satisfaction class containing a given sentence $\varphi$. In particular we show that there exists a full satisfaction class containing the negation of a nonstandard disjunction of copies of the sentence $0=0$. Lachlan has proved the converse of our main result for countable models, that is, a countable model of PA for which there is a full satisfaction class is necessarily recursively saturated. His theorem is presented in the paper immediately following this one. It is an open question whether there is a recursively saturated uncountable model of PA which has no full satisfaction class. The referee has suggested that the model of Kaufmann [4] might be an example.

It will be clear that our results apply not only to arithmetic but also to any theory in which a sufficient part of arithmetic can be interpreted and which permits the coding of finite sequences of individuals. The theorem for ZF corresponding to our main result is much easier, see Proposition 5.1 of [5].

The plan of the paper is as follows. In the first section we define the particular system of logic we need and establish some lemmas about it. In the second section we show that given a proof of finite height we can replace each formula by a standard approximation in such a way that the proof remains a proof. In the third and final section we derive the main theorem and related results.

We are grateful to J. Paris for suggesting the use of $\omega$-logic for the construction of satisfaction classes and to the referee whose many suggestions regarding an earlier version have clarified many parts of the paper.

**1. $\mathcal{M}$-logic.** As above $\mathcal{M}$ is a model of PA. Our logic will be denoted $\mathcal{L}(\mathcal{M})$ and its formulas are all sentences of $^*L(\mathcal{M})$. The axioms of $\mathcal{L}(\mathcal{M})$ are all formulas of the form $\neg\varphi\vee\varphi$ together with all true atomic and negated atomic sentences. The rules of inference of $\mathcal{L}(\mathcal{M})$ are

Expansion $\dfrac{\psi}{\varphi\vee\psi}$

Commutativity $\dfrac{\varphi\vee\psi}{\psi\vee\varphi}$

Associativity $\dfrac{\varphi\vee(\psi\vee\chi)}{(\varphi\vee\psi)\vee\chi}$ $\dfrac{(\varphi\vee\psi)\vee\chi}{\varphi\vee(\psi\vee\chi)}$

Contraction $\dfrac{\varphi\vee\varphi}{\varphi}$

Cut $\dfrac{\chi \vee \varphi \quad \neg\chi \vee \psi}{\varphi \vee \psi}$

$\exists$-introduction $\dfrac{\varphi(a)}{\exists x \varphi(x)}$ ($a$ in $\mathcal{M}$)

$\mathcal{M}$-rule $\dfrac{\neg\varphi(a) \text{ for all } a \text{ in } \mathcal{M}}{\neg\exists x \varphi(x)}$

and each rule is repeated with the addition of an arbitrary formula $\theta$ as a parameter. For example the repetition of the $\mathcal{M}$-rule is:

$$\frac{\theta \vee \neg\varphi(a) \text{ for all } a \text{ in } \mathcal{M}}{\theta \vee \neg\exists x \varphi(x)}.$$

We use the symbol $\vdash$ to denote provability in $\mathcal{L}(\mathcal{M})$. To make this notion precise, for any set of sentences $\Gamma \subseteq {}^*L(\mathcal{M})$ let $\Gamma'$ denote the set

$\Gamma \cup \{\theta : \theta \in {}^*L(\mathcal{M})$ & $\theta$ is an immediate consequence of a sentence or sentences in $\Gamma$ by one of the rules$\}$.

Define $\Gamma^\alpha$ by induction on the ordinal $\alpha$:

$$\Gamma^\alpha = \Gamma \cup \bigcup_{\beta<\alpha} (\Gamma^\beta)'.$$

Then $\Gamma \vdash \theta$ means that there exists $\alpha \in On$ such that $\theta \in \Gamma^\alpha$. Because of the $\mathcal{M}$-rule a proof of $\theta$ from $\Gamma$ may be infinite and may involve an infinite external subset of $\Gamma$.

The reason for repeating the rules with a parameter is that it allows an easy proof of:

LEMMA 1. *Let $\Gamma \subseteq {}^*L(\mathcal{M})$ be a set of sentences and $\varphi, \psi \in {}^*L(\mathcal{M})$ be sentences, then*

$$\Gamma \bigcup \{\varphi\} \vdash \psi \quad \textit{implies} \quad \Gamma \vdash \neg\varphi \vee \psi$$

*and*

$$\Gamma \bigcup \{\varphi\} \vdash \psi, \neg\psi \quad \textit{implies} \quad \Gamma \vdash \neg\varphi.$$

**Proof.** The second half of the conclusion follows easily from the first using the cut rule. For the first half suppose a proof of $\psi$ from $\Gamma \cup \{\varphi\}$ is given, then the idea is that we go through the proof replacing each formula $\alpha$ by $\neg\varphi \vee \alpha$ so that occurrences of $\varphi$ now become axioms. If $\alpha$ is an axiom or in $\Gamma$ then we have $\Gamma \vdash \neg\varphi \vee \alpha$ by means of the first two rules. It remains to check that if we repeat the process of adding a parameter to the rules of inference then the new rules we get are already derivable. We leave this to the reader.

A set of sentences $\Gamma \subseteq {}^*L(\mathcal{M})$ is called *consistent* if no contradiction is derivable from it in $L(\mathcal{M})$.

LEMMA 2. *Let* $\Gamma \subseteq {}^*L(\mathcal{M})$ *be a set of sentences and* $\mathcal{M}$ *be countable.* $\Gamma$ *can be extended to a full satisfaction class iff* $\Gamma$ *is consistent.*

**Proof.** Suppose $\Gamma$ is consistent. Let $\langle \varphi_i : i < \omega \rangle$ be an enumeration of the sentences in ${}^*L(\mathcal{M})$. We define an increasing sequence $\langle \Gamma_i : i < \omega \rangle$ of consistent subsets of ${}^*L(\mathcal{M})$ such that $\Gamma_0 = \Gamma$ and for all $i < \omega$, either $\varphi_i$ or $\neg \varphi_i$ is in $\Gamma_i$. Further for each $i$, if $\varphi_i$ has the form $\exists x \psi(x)$ and is in $\Gamma_i$, then $\psi(a) \in \Gamma_i$ for some $a \in \mathcal{M}$. From Lemma 1 it is easy to deduce that such a sequence $\langle \Gamma_i : i < \omega \rangle$ exists. Clearly $\cup \{\Gamma_i : i < \omega\}$ is a full satisfaction class. The proof in the other direction is immediate.

If the model $\mathcal{M}$ is standard it is clear that the empty set is consistent in $\mathcal{L}(\mathcal{M})$. However, for a nonstandard model the consistency of $\emptyset$ is not obvious because the nonstandard sentences have no natural interpretation and indeed the main result of this paper together with Lachlan's converse show that, for countable nonstandard $\mathcal{M}$, $\emptyset$ is consistent in $\mathcal{L}(\mathcal{M})$ iff $\mathcal{M}$ is recursively saturated.

**2. Approximation of nonstandard formulas.** It is convenient to introduce a new logic $\mathcal{L}'(\mathcal{M})$. Let $L'$ be the language obtained by adjoining to $L$ infinitely many predicate symbols of each finite arity. Let $L'(\mathcal{M})$ be obtained by adjoining to $L'$ all the variables in the sense of $\mathcal{M}$ and constants naming all the elements of $\mathcal{M}$. Formulas are built up in the usual manner using the same logical operators as before but here there are only standard formulas. As axioms we take all sentences of the form $\neg \varphi \vee \varphi$ and all atomic and negated atomic sentences which are true in $\mathcal{M}$. The rules of inference are the same as for $\mathcal{L}(\mathcal{M})$; again only sentences are admitted to proof trees. The symbol $\vdash$ will also be used to denote provability in $\mathcal{L}'(\mathcal{M})$.

Let $\psi[p_0, \ldots, p_{k-1}]$ be a sentence of $L'(\mathcal{M})$ where all the new predicate symbols are displayed, and where $p_i$ is $n_i$-ary. Let $\varphi_0(x_1^0, \ldots, x_{n_0}^0), \ldots, \varphi_{k-1}(x_1^{k-1}, \ldots, x_{n_{k-1}}^{k-1})$ be formulas either all in $L'(\mathcal{M})$ or all in ${}^*L(\mathcal{M})$, where all free variables are displayed. We use $\psi[\varphi_0/p_0, \ldots, \varphi_{k-1}/p_{k-1}]$ to denote the formula obtained by replacing each part of $\psi$ of the form $p_i(t_1, \ldots, t_{n_i})$ by $\varphi_i(t_1, \ldots, t_{n_i})$. If $\varphi$ in $L'(\mathcal{M}) \cup {}^*L(\mathcal{M})$ can be written as $\psi[\varphi_0/p_0, \ldots, \varphi_{k-1}/p_{k-1}]$ then we call $\psi$ an *approximation of* $\varphi$.

Let $L(\mathcal{M}, p)$ be the language obtained by adjoining a new propositional constant $p$ to $L(\mathcal{M})$. Let $\varphi \in {}^*L(\mathcal{M})$. By a *part of* $\varphi$ we mean a pair $\langle \varphi', \psi \rangle$ where $\varphi' \in {}^*L(\mathcal{M}, p)$, $\psi \in {}^*L(\mathcal{M})$, $p$ occurs exactly once in $\varphi'$, and $\varphi$ results from $\varphi'$ by replacing $p$ by $\psi$. The set of all parts of $\varphi$ will be denoted by $\Pi(\varphi)$.

The depth of the *part* $\langle \varphi', \psi \rangle$, denoted $d(\varphi', \psi)$, is the number of logical operators of $\varphi'$ within whose scope $p$ lies. The *depth* of a formula is the maximum depth of its parts.

We call $\psi$ the *associated formula* of the part $\langle \varphi', \psi \rangle$. Let $\langle \varphi_0', \psi_0 \rangle$ and $\langle \varphi_1', \psi_1 \rangle$ be parts of $\varphi$ then $\langle \varphi_0', \psi_0 \rangle$ *includes* $\langle \varphi_1', \psi_1 \rangle$ written $\langle \varphi_0', \psi_0 \rangle \succ \langle \varphi_1', \psi_1 \rangle$ if $\varphi_1'$ can be obtained from $\varphi_0'$ by replacing $p$ by some formula of $^*L(\mathcal{M}, p)$. Note that $>$ is transitive.

Let $L(\mathcal{M}, z)$ be the language obtained by adjoining a new individual variable $z$ to $L(\mathcal{M})$. Let $\varphi \in {}^*L(\mathcal{M})$. By an *occurrence in* $\varphi$ we mean a pair $\langle \varphi', t\rangle$ where $\varphi' \in {}^*L(\mathcal{M}, z)$, $t$ is a term of $^*L(\mathcal{M})$, $z$ occurs exactly once in $\varphi'$ and not in a quantifier, $\varphi$ results from $\varphi'$ by substituting $t$ for $z$, and $t$ is free for $z$ in $\varphi'$. It should be clear that this definition simply makes precise the usual notion of a free occurrence of a term in a formula. Let the set of occurrences in $\varphi$ be denoted $\mathcal{O}(\varphi)$. Let $\bar{\varphi}$ denote the formula of $^*L(\mathcal{M}, z)$ obtained by replacing every occurrence in $\varphi$ by $z$.

Let the concepts of part and occurrence together with the related terminology and notation be carried over to other languages in the obvious way.

Let $^*T(\mathcal{M})$ denote the set of terms of $^*L(\mathcal{M})$. Every formula $\varphi$ in $^*L(\mathcal{M})$ has associated with it a unique mapping $t_\varphi : \mathcal{O}(\bar{\varphi}) \to {}^*T(\mathcal{M})$ such that, if $o$ is replaced by $t_\varphi(o)$, for every occurrence $o$ in $\bar{\varphi}$ simultaneously, then $\varphi$ results.

Call two formulas $\varphi_0$, $\varphi_1$ in $^*L(\mathcal{M})$ *weakly equivalent*, written $\varphi_0 \sim \varphi_1$, if $\bar{\varphi}_0 = \bar{\varphi}_1$. Call $\varphi_0, \varphi_1$ *equivalent*, written $\varphi_0 \approx \varphi_1$, if $\varphi_0 \sim \varphi_1$ and there exists an equivalence relation $E$ on $\mathcal{O}(\bar{\varphi}_0) = \mathcal{O}(\bar{\varphi}_1)$ such that $t_{\varphi_0}$, $t_{\varphi_1}$ are well defined on $\mathcal{O}(\bar{\varphi}_0)/E$ and differ on at most a finite number of equivalence classes. There is a canonical candidate for $E$ given by:

$$o_0 E o_1 \Leftrightarrow [t_{\varphi_0}(o_0) = t_{\varphi_0}(o_1) \ \&\ t_{\psi_1}(o_0) = t_{\psi_1}(o_1)].$$

It is easy to show that $\approx$ is an equivalence relation. Further, suppose $\Phi$ is a finite set of pairwise strongly equivalent formulas each having only a finite number of free variables. We can associate with $\Phi$ a certain formula $\theta_\Phi$ as follows. Let $\mathcal{O}(\Phi)$ be the common value of $\mathcal{O}(\bar{\varphi})$ for all $\varphi \in \Phi$. Define an equivalence relation $E_\Phi$ on $\mathcal{O}(\Phi)$ by:

$$o_0 E_\Phi o_1 \Leftrightarrow \bigwedge\{t_\varphi(o_0) = t_\varphi(o_1) : \varphi \in \Phi\}.$$

For $\varphi \in \Phi$ the map of $\mathcal{O}(\Phi)/E_\Phi$ into $^*T(\mathcal{M})$ induced by $t_\varphi$ will also be denoted $t_\varphi$. If $\varphi$, $\psi \in \Phi$, then $\varphi \approx \psi$ and thus there are at most a finite number of classes in $\mathcal{O}(\Phi)/E_\Phi$ on which $t_\varphi$, $t_\psi$ disagree. Let $\mathcal{C}_1, \ldots, \mathcal{C}_l$ be an enumeration of all classes $\mathcal{C}$ in $\mathcal{O}(\Phi)/E_\Phi$ such that there exist $\varphi$, $\psi \in \Phi$ such that $t_\varphi(\mathcal{C}) \neq t_\psi(\mathcal{C})$ or $t_\varphi(\mathcal{C})$ is a variable. Now let $z_1, \ldots, z_l$ be new variables. Fix $\varphi \in \Phi$. In $\bar{\varphi}$ replace each occurrence $o$ in $\mathcal{C}_i$, $1 \leq i \leq l$, by $z_i$ and every occurrence $o$ not in any $\mathcal{C}_i$ by $t_\varphi(o)$. Let the resulting formula be denoted by $\theta_\Phi(z_1, \ldots, z_l)$ and notice that this formula is unique to within a permutation of the variables. That is, there is no dependence on the choice of $\varphi$. Notice also that for each $\varphi \in \Phi$ there exist unique $t_1, \ldots, t_l$ such that $\varphi$ is $\theta_\Phi(t_1, \ldots, t_l)$.

We are now ready to begin the series of definitions which will lead directly to the crucial concept of the paper that of the $n$th approximation of a formula of $^*L(\mathcal{M})$.

Fix $n<\omega$ and a sentence $\varphi\in{}^*L(\mathcal{M})$. For $m<\omega$ we define $\Pi^{(m)}(\varphi, n)$ by induction as follows:

$$\Pi^{(0)}(\varphi, n)=\{\langle\varphi', \psi\rangle\in\Pi(\varphi): d(\varphi', \psi)\leq n\}$$

$$\Pi^{(m+1)}(\varphi, n)=\{\langle\varphi', \psi\rangle\in\Pi(\varphi): \exists\varphi_0'\exists\psi_0\exists\varphi_1'\exists\psi_1[\langle\varphi_0', \psi_0\rangle\in\Pi^{(m)}(\varphi, n)$$

$$\& \langle\varphi_1', \psi_1\rangle\in\Pi^{(0)}(\varphi, n) \& \psi_0\approx\psi_1 \& \langle\varphi', \psi\rangle\prec\langle\varphi_0', \psi_0\rangle$$

$$\& d(\varphi_1', \psi_1)+d(\varphi', \psi)\leq d(\varphi_0', \psi_0)+n]\}.$$

For all sufficiently large $m<\omega$, $\Pi^{(m)}(\varphi, n)$ is fixed. To see this notice that for each $m<\omega$, if $\langle\varphi', \psi\rangle\in\Pi^{(m)}(\varphi, n)$, then there exists $\langle\varphi_1', \psi_1\rangle\in\Pi^{(0)}(\varphi, n)$ such that $\psi\approx\psi_1$. Also, if $\langle\varphi_0', \psi_0\rangle\in\Pi(\varphi)$, $\langle\varphi_1', \psi_1\rangle\in\Pi^{(m)}(\varphi, n)$, and $\langle\varphi_0', \psi_0\rangle\succ\langle\varphi_1', \psi_1'\rangle$, then $\langle\varphi_0', \psi_0\rangle\in\Pi^{(m)}(\varphi, n)$. Thus, if $\langle\varphi', \psi\rangle\in\Pi^{(m)}(\varphi, n)$, then there exist $\langle\varphi_i', \psi_i\rangle\in\Pi^{(m)}(\varphi, n)$, $i\leq d(\varphi', \psi)$, such that

$$\langle\varphi_i', \psi_i\rangle\gneq\langle\varphi_{i+1}', \psi_{i+1}\rangle \qquad (i<d(\varphi', \psi))$$

and $\langle\varphi_i', \psi_i\rangle=\langle\varphi', \psi\rangle$ for $i=d(\varphi', \psi)$. Now the crucial point is that $\psi_i\not\approx\psi_j$ for $i<j\leq d(\varphi', \psi)$. Also $|\Pi^{(0)}(\varphi, n)|\leq 2^n$ and as noted above each $\psi_i$ is equivalent to a member of $\Pi^{(0)}(\varphi, n)$. Thus

$$\langle\varphi', \psi\rangle\in\Pi^{(m)}(\varphi, n)\Rightarrow d(\varphi', \psi)\leq 2^n.$$

Since $\Pi^{(m)}(\varphi, n)\subset\Pi^{(m+1)}(\varphi, n)$ it is clear that there exists $k<\omega$ such that

$$k\leq m<\omega\Rightarrow\Pi^{(m)}(\varphi, n)=\Pi^{(k)}(\varphi, n).$$

For such $k$ let

$$\Gamma(\varphi, n)=\{\psi\in{}^*L(\mathcal{M}): \exists\varphi'(\langle\varphi', \psi\rangle\in\Pi^{(k)}(\varphi, n))\}$$

and

$$\Gamma_I(\varphi, n)=\{\psi\in{}^*L(\mathcal{M}): \exists\varphi'(\langle\varphi', \psi\rangle\in\Pi^{(k)}(\varphi, n)$$

and $\langle\varphi', \psi\rangle$ is $\prec$-minimal in $\Pi^{(k)}(\varphi, n)\}$.

Intuitively $\Gamma(\varphi, n)$ is the least set of subformulas of $\varphi$ including all those of depth $\leq n$ in $\varphi$ such that, if $\psi_0\in\Gamma(\varphi, n)$ is equivalent to $\psi_1\in\Gamma(\varphi, n)$ of depth $\leq n$ in $\varphi$, then $\Gamma(\varphi, n)$ contains a subformula of $\psi_0$ if and only if it contains the corresponding subformula of $\psi_1$. Also $\Gamma_I(\varphi, n)$ consists of the members of $\Gamma(\varphi, n)$ which are minimal in $\Gamma(\varphi, n)$. However, a precise treatment of these concepts requires the consideration of parts of $\varphi$ rather than just subformulas, because strictly speaking one cannot speak of the depth of a subformula. Further, if $\psi_0\approx\psi_1$, there is a natural bijection of $\Pi(\psi_0)$ onto $\Pi(\psi_1)$ which underlies the construction of $\Gamma(\varphi, n)$ and which has no analogue if we speak only of subformulas.

From above every $\psi\in\Gamma(\varphi, n)$ is the associated formula of a part of $\varphi$ of depth $\leq 2^n$, whence $|\Gamma(\varphi, n)|\leq 2^{2^n}$. Let $\Gamma_I(\varphi, n)/\approx$ be the set of equivalence classes into which $\Gamma_I(\varphi, n)$ is partitioned by $\approx$. For $\Phi\in\Gamma_I(\varphi, n)/\approx$ we let

https://doi.org/10.4153/CMB-1981-045-3 Published online by Cambridge University Press

$\theta_\Phi(z_1, \ldots, z_l)$ be the canonical formula associated with $\Phi$ described above. Call $\Phi \in \Gamma_I(\varphi, n)/\approx$ *atomic* if it contains an atomic formula in which case every member of $\Phi$ is atomic.

We now define a mapping $F_{\varphi,n}$ of $\Gamma(\varphi, n)$ into $L'(\mathcal{M})$. If $\psi \in \Gamma(\varphi, n)$ is atomic, let $F_{\varphi,n}(\psi) = \psi$. For each nonatomic $\Phi \in \Gamma_I(\varphi, n)/\approx$ with associated formula $\theta_\Phi(z_1, \ldots, z_l)$ choose a distinct new $l$-ary predicate symbol $p_\Phi$ of $L'$ and for each $\psi = \theta_\Phi(t_1, \ldots, t_l) \in \Phi$ let $F_{\varphi,n}(\psi) = p_\Phi(t_1, \ldots, t_l)$. Extend $F_{\varphi,n}$ to the rest of $\Gamma(\varphi, n)$ by the rules:

$$F_{\varphi,n}(\psi_0 \vee \psi_1) = F_{\varphi,n}(\psi_0) \vee F_{\varphi,n}(\psi_1)$$
$$F_{\varphi,n}(\neg\psi) = \neg F_{\varphi,n}(\psi)$$
$$F_{\varphi,n}(\exists x\psi) = \exists x F_{\varphi,n}(\psi).$$

We call $F_{\varphi,n}(\varphi)$ *the n-th approximation of* $\varphi$. There is an element of nonuniqueness in the definition of $n$th approximation but it is inessential.

As an example consider a sentence $\varphi$ in $^*L(\mathcal{M})$ of the form

$$\exists x(\theta(x, a) \vee \psi(x)) \vee \neg\exists y(\theta(b, y) \vee \psi(y))$$

where $\theta(x, a) \not\approx \psi(x)$. The zeroth, first, and second approximations of $\varphi$ are $p$, $p_0 \vee p_1$, $\exists x p_0(x) \vee \neg p_1$ respectively and in each case the number $k$ of the above definition is zero. However, since $\theta(x, a) \vee \psi(x) \approx \theta(b, y) \vee \psi(y)$, when the third approximation of $\varphi$ is constructed we shall have $k = 1$ and the approximation will be

$$\exists x(p_0(x, a) \vee p_1(x)) \vee \neg\exists y(p_0(b, y) \vee p_1(y)).$$

Let $\langle \varphi^0, \ldots, \varphi^{k-1} \rangle = \boldsymbol{\varphi}$ be a finite sequence of sentences of $^*L(\mathcal{M})$ with $k$ standard and let $n < \omega$ be fixed as before. We define the *n-th approximation of* $\boldsymbol{\varphi}$, which will be a sequence of sentences of $L'(\mathcal{M})$ of length $k$, by carrying out the above construction simultaneously for all members of $\boldsymbol{\varphi}$. Specifically, we define $\Pi^{(m)}(\boldsymbol{\varphi}, n)$ for $m < \omega$ by:

$$\Pi^{(0)}(\boldsymbol{\varphi}, n) = \cup\{\Pi^{(0)}(\varphi^i, n) : i < k\},$$

$$\Pi^{(m+1)}(\boldsymbol{\varphi}, n) = \{\langle \varphi', \psi \rangle \in \bigcup_{i<k} \Pi(\varphi^i) : \exists i \exists \varphi'_0 \exists \psi_0 \exists j \exists \varphi'_1 \exists \psi_1 [i, j < k \;\&$$
$$\langle \varphi'_0, \psi_0 \rangle \in \Pi^{(m)}(\varphi^i, n) \;\&\; \langle \varphi'_1, \psi_1 \rangle \in \Pi^{(0)}(\varphi^j, n) \;\&\; \psi_0 \approx \psi_1 \;\&\; \langle \varphi', \psi \rangle$$
$$\prec \langle \varphi'_0, \psi_0 \rangle \;\&\; d(\varphi'_1, \psi_1) + d(\varphi', \psi) \leq d(\varphi'_0, \psi_0) + n]\}.$$

Reasoning in the same way as before we see that each $\langle \varphi', \psi \rangle$ in $\Pi^{(m)}(\boldsymbol{\varphi}, n)$ has depth $\leq k \cdot 2^n$ and that there exists $l$ such that

$$l \leq m < \omega \Rightarrow \Pi^{(m)}(\boldsymbol{\varphi}, n) = \Pi^{(l)}(\boldsymbol{\varphi}, n).$$

For such $l$ let

$$\Gamma(\varphi, n) = \{\psi \in L(\mathcal{M}) : \exists \varphi'(\langle \varphi', \psi \rangle \in \Pi^{(l)}(\boldsymbol{\varphi}, n))\}$$

https://doi.org/10.4153/CMB-1981-045-3 Published online by Cambridge University Press

and let $\Gamma_I(\boldsymbol{\varphi}, n)$ consist of the formulas in $\Gamma(\boldsymbol{\varphi}, n)$ which are $\prec$-minimal. Now we define $F_{\boldsymbol{\varphi},n}$ on $\Gamma(\boldsymbol{\varphi}, n)$ in a manner similar to that in which we defined $F_{\varphi,n}$ on $\Gamma(\varphi, n)$, and the $n$th approximation of $\boldsymbol{\varphi}$ is the sequence $\langle F_{\boldsymbol{\varphi},n}(\varphi^0), \ldots, F_{\boldsymbol{\varphi},n}(\varphi^{k-1})\rangle$.

LEMMA 3. *Let $\varphi$ be a sentence of $^*L(\mathcal{M})$ and for $i<\omega$ let $\varphi_i$ denote the $i$-th approximation of $\varphi$. Then $\varphi_i$ is an approximation of $\varphi$ and if $i<j<\omega$ then $\varphi_i$ is an approximation of $\varphi_j$.*

LEMMA 4. *Let $\psi$, $\chi$ be sentences of $L'(\mathcal{M})$ and $\psi$ be an approximation of $\chi$. If there is a proof in $\mathcal{L}'(\mathcal{M})$ from $\emptyset$ of $\psi$ of height $n$, then the same is true of $\chi$.*

LEMMA 5. *Let $\varphi$ be a sentence of $^*L(\mathcal{M})$ and $\psi$ be an approximation of $\varphi$. If $\psi$ has depth $n$ then $\psi$ is an approximation of the $n$-th approximation of $\varphi$.*

LEMMA 6. *Let $\theta \vee \neg\exists x\varphi(x)$ be a sentence of $^*L(\mathcal{M})$, and $\theta' \vee \neg\exists x\varphi'(x)$ be its $(k+2)$-th approximation, then the $k$-th approximation of $\theta \vee \neg\varphi(a)$ is an approximation of $\theta' \vee \neg\varphi'(a)$.*

The proof of these lemmas is a tedious exercise which we omit. The principal result of this section is

LEMMA 7. *There exists a recursive function $G$ such that for every sentence $\varphi$ of $^*L(\mathcal{M})$ if there is a proof of $\varphi$ from $\emptyset$ in $\mathcal{L}(\mathcal{M})$ of height $n<\omega$, then there is a proof in $\mathcal{L}'(\mathcal{M})$ from $\emptyset$ of height $n$ of the $G(n)$-th approximation to $\varphi$.*

**Proof.** As before $n$ denotes a standard positive integer. The second approximation of any sentence $\neg\psi \vee \psi$ has the same form, and any approximation of an atomic sentence is the sentence itself. Thus we can take $G(1)=2$. For the rest we proceed by induction on $n$. It is clearly sufficient to establish an appropriate sublemma for each rule of inference. We shall show how to handle three of the rules and leave the other cases to the reader.

SUBLEMMA (First associative rule). *Let $k$, $n<\omega$ and $\varphi$, $\chi$, $\psi$ be sentences of $^*L(\mathcal{M})$. If there is a proof in $\mathcal{L}'(\mathcal{M})$ from $\emptyset$ of height $n$ of the $k$-th approximation of $\varphi \vee (\psi \vee \chi)$ then there is a proof in $\mathcal{L}'(\mathcal{M})$ from $\emptyset$ of height $n+1$ of the $(k+2)$-th approximation of $(\varphi \vee \psi) \vee \chi$.*

**Proof of sublemma.** Let $\varphi_k \vee (\psi_k \vee \chi_k)$ and $(\varphi_{k+2} \vee \psi_{k+2}) \vee \chi_{k+2}$ be the $k$th and $(k+2)$th approximations of $\varphi \vee (\psi \vee \chi)$ and $(\varphi \vee \psi) \vee \chi$ respectively. In the first step of the construction of $(\varphi_{k+2} \vee \psi_{k+2}) \vee \chi_{k+2}$ we explore $\varphi$, $\psi$ to depth $k$ and $\chi$ to depth $k+1$, whereas in the first step of the construction of $\varphi_k \vee (\psi_k \vee \chi_k)$ we explore $\varphi$ to depth $k-1$ and $\psi$, $\chi$ to depth $k-2$. Because the analysis of $(\varphi \vee \psi) \vee \chi$ goes deeper than that of $\varphi \vee (\psi \vee \chi)$, $(\varphi_k \vee \psi_k) \vee \chi_k$ is an approximation of $(\varphi_{k+2} \vee \psi_{k+2}) \vee \chi_{k+2}$. The reason why $k+2$ is required here rather than $k+1$ is illustrated by the formulas:

$$((\theta_0 \vee \theta_1) \vee \theta_1) \vee (\theta_0 \vee \theta_1), \; (((\theta_0 \vee \theta_1) \vee \theta_1) \vee \theta_0) \vee \theta_1)$$

where $\theta_0$, $\theta_1$ are inequivalent atomic sentences. The second approximation of the first is

$$((p_0 \vee p_1) \vee p_1) \vee (p_0 \vee p_1)$$

but to get this deep an analysis of the second sentence we have to take its fourth approximation.

SUBLEMMA (Cut rule). *Let $k$, $n < \omega$ and $\varphi$, $\psi$, $\chi$ be sentences of $^*L(\mathcal{M})$. If there are proofs in $\mathcal{L}'(\mathcal{M})$, from $\emptyset$ of height $n$, of the $k$-th approximation of $\chi \vee \varphi$ and $\neg\chi \vee \psi$, then there is a proof in $\mathcal{L}'(\mathcal{M})$, from $\emptyset$ of height $n+1$, of the $3 \cdot 2^k$-th approximation of $\varphi \vee \psi$.*

**Proof of sublemma.** Let $(\chi \vee \varphi)_k$, $(\neg\chi \vee \psi)_k$ denote the $k$th approximations of $\chi \vee \varphi$, $\neg\chi \vee \psi$ respectively. Let $\langle \chi_k, \varphi_k, \psi_k \rangle$ be the $k$th approximation of $\langle \chi, \varphi, \psi \rangle$. Now $(\chi \vee \varphi)_k$ is an approximation of $\chi_k \vee \varphi_k$ whence by Lemma 4 there is a proof of $\chi_k \vee \varphi_k$ of height $n$. Similarly there is a proof of $\neg\chi_k \vee \psi_k$ of height $n$. Thus by the cut rule there is a proof of $\varphi_k \vee \psi_k$ of height $n+1$. The depth of $\varphi_k \vee \psi_k$ is $\leq 3 \cdot 2^k$ whence, letting $(\varphi \vee \psi)_k$ denote the $3 \cdot 2^k$th approximation of $\varphi \vee \psi$, by Lemma 5 $\varphi_k \vee \psi_k$ is an approximation of $(\varphi \vee \psi)_k$. The conclusion follows by Lemma 4.

SUBLEMMA ($\mathcal{M}$-rule with parameter). *Let $k$, $m$, $n < \omega$ and $\theta \vee \neg\exists x\varphi(x)$ be a formula of $^*L(\mathcal{M})$. Suppose there is a proof in $\mathcal{L}'(\mathcal{M})$, from $\emptyset$ of height $<n$, of the $k$-th approximation of $\theta \vee \neg\varphi(a)$ for every $a \in \mathcal{M}$. Then there is a proof in $\mathcal{L}'(\mathcal{M})$, from $\emptyset$ of height $\leq n$, of the $(k+2)$-th approximation of $\theta \vee \neg\exists x\varphi(x)$.*

**Proof.** Let $\theta' \vee \neg\exists x\varphi'(x)$ be the $(k+2)$th approximation of $\theta \vee \neg\exists x\varphi(x)$, then by Lemma 6 the $k$th approximation of $\theta \vee \neg\varphi(a)$ is an approximation of $\theta' \vee \neg\varphi'(a)$. By Lemma 4 for every $a \in \mathcal{M}$ there is a proof in $\mathcal{L}'(\mathcal{M})$, from $\emptyset$ of height $<n$, of $\theta' \vee \neg\varphi'(a)$. The conclusion now follows by applying the $\mathcal{M}$-rule.

**3. The main result.** We continue the conventions of previous sections. A crucial observation is the following:

LEMMA 8. *If $\mathcal{M}$ is recursively saturated, $\Gamma \cup \{\varphi\} \subseteq {}^*L(\mathcal{M})$ is a set of sentences definable in $\mathcal{M}$ with parameters, and $\Gamma \vdash \varphi$, then there is a proof of $\varphi$ from $\Gamma$ of finite height.*

We omit the proof because this follows at once from the fact that over a recursively saturated model all first-order inductive definitions close off by stage $\omega$. See [1], Corollary V1.5.13, or [7], Exercise 4.7.

Let $Th(\mathcal{M})$ denote the set of all standard sentences of $L(\mathcal{M})$ which are true in $\mathcal{M}$. Let $\vdash_{PC}$ denote derivability in the predicate calculus.

LEMMA 9. *Let $\mathcal{M}$ be resplendent and $\psi$ be a sentence of $L'(\mathcal{M})$. Then $\vdash \psi$ if and only if $Th(\mathcal{M}) \vdash_{PC} \psi$.*

**Proof.** Observe that the logic $\mathcal{L}'(\mathcal{M})$ was designed so that $\vdash \psi$ if and only if $\psi$ is true in $\mathcal{M}$ under every possible interpretation of the new predicate symbols. Thus the "if" part of the conclusion is clear. For the other direction suppose $Th(\mathcal{M}) \nvdash_{PC} \psi$. Then there exists an elementary extension $\mathcal{M}^*$ of $\mathcal{M}$ and interpretations of the new predicate symbols of $\psi$ over $\mathcal{M}^*$ which make $\psi$ false. Since $\mathcal{M}$ is resplendent such interpretations already exist over $\mathcal{M}$, whence $\nvdash \psi$.

THEOREM. *Let $\mathcal{M}$ be a countable recursively saturated model of PA. There exists a full satisfaction class $\Sigma$ for $\mathcal{M}$. Moreover, for any sentence $\varphi \in {}^*L(\mathcal{M})$, $\Sigma$ can be found containing $\varphi$ if and only if there is no approximation $\psi$ of $\varphi$ such that $Th(\mathcal{M}) \vdash_{PC} \neg \psi$.*

**Proof.** It is obviously sufficient to prove the second part of the theorem. If $Th(\mathcal{M}) \vdash_{PC} \neg \psi$ for some approximation $\psi$ of $\varphi$ then it is immediate from the definitions that there is no full satisfaction class containing $\varphi$. Suppose there is no such approximation $\psi$. By Lemma 9 there is no approximation $\psi$ of $\varphi$ such that $\vdash \neg \psi$. From Lemma 2 we have the desired conclusion provided $\{\varphi\}$ is consistent. If not, then $\vdash \neg \varphi$ and by Lemma 8 there is a proof of $\neg \varphi$ from $\emptyset$ of finite height. By Lemma 7 we have $\vdash \neg \psi$ for some approximation $\psi$ of $\varphi$. This contradiction shows that $\{\varphi\}$ is consistent and completes the proof of the theorem.

Using Lachlan's converse the theorem can be strengthened to say that a countable model of PA is recursively saturated if and only if it has a full satisfaction class. Also, using $\Pi_1^1$-reflection (see Theorem 2.4 (vi) of [2]) we can replace "countable recursively saturated" in the statement of the theorem by "resplendent". By a result of Harnik (see Theorem 9.3 of [6]) for $\mathcal{M}$ countable there will be $2^{\aleph_0}$ possibilities for $\Sigma$. Other results about the number of various satisfaction classes are contained in [5].

As an example consider the sequence of sentences of $L(\mathcal{M})$ defined as follows: $\theta_0$ is $0=0$, and $\theta_{i+1} = (\theta_i \vee \theta_i)$ for all $i$. For each standard $i$, $\theta_i$ obviously belongs to every full satisfaction class. However, suppose $a \in \mathcal{M}$ is nonstandard then the successive approximations of $\theta_a$ are: $p$, $p \vee p$, $(p \vee p) \vee (p \vee p), \ldots$ where $p$ is a propositional constant. Since none of these approximations is a logical consequence of $Th(\mathcal{M})$, no approximation of $\theta_a$ is a logical consequence of $Th(\mathcal{M})$. Hence $\{\neg \theta_a\}$ is consistent and there is a full satisfaction class containing $\neg \theta_a$.

Our theorem can be somewhat sharpened as follows: Let $\Theta$ be an $L(\mathcal{M})$-axiom schema all standard closed instances of which are true in $\mathcal{M}$. Note that by "standard instance" here we mean that the length of the formula is standard so that its meaning is unequivocal; if there are free variables we allow arbitrary substitution for them. As an example one can take the axiom schema of induction. Let $\Theta$ also denote the set of all closed instances of the given axiom schema in the sense of $\mathcal{M}$. Then $\Theta \subseteq L(\mathcal{M})$. Using the same technique as above and with the same assumptions we can show that there is a full satisfaction class $\Sigma \supseteq \Theta$. Further for any $\varphi \in L(\mathcal{M})$, $\Sigma$ can be found containing $\varphi$ if and only if there is no approximation $\psi$ of $\varphi$ such that $\mathrm{Th}(\mathcal{M}) \cup E \vdash_{PC} \neg \psi$ where $E$ denotes the set of all closed instances of the given axiom schema in $L'(\mathcal{M})$. Thus there are full satisfaction classes making some $\theta_a$ false but all instances, in the sense of $\mathcal{M}$, of the induction schema true.

## References


1. K. J. Barwise. *Admissible Sets and Structures*, Springer, Berlin 1975.
2. K. J. Barwise and J. Schlipf. *An introduction to recursively saturated and resplendent models*, J. Symb. Logic **41** (1976), 531–536.
3. J. Geiser. *Nonstandard logic*, J. Symb. Logic **33** (1968), 236–250.
4. M. Kaufmann. *A rather classless model*, Proc. American Math. Soc. **62** (1977), 330–333.
5. S. Krajewski. *Non-standard satisfaction classes*, in Set theory and Hierarchy theory, Springer Lecture Notes No. 537, 1976, 121–145.
6. M. Makkai. *Admissible sets and infinitary logic*, Handbook of Mathematical Logic, North-Holland, Amsterdam 1977, 233–281.
7. Y. N. Moschovakis. *Elementary Induction on Abstract Structures*, North-Holland, Amsterdam 1974.
8. A. Robinson. *On languages based on non-standard arithmetic*, Nagoya Math. J. **22** (1963), 83–107.



Simon Fraser University
Burnaby 2, B.C. V5A 1S6